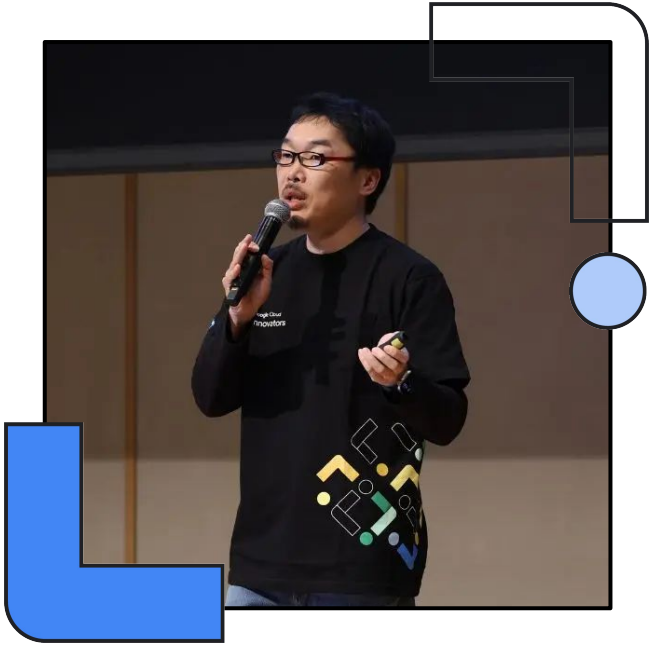
Google Cloud

生成 AI が拓くデータ活用の新境地

Google Cloud の「データ エージェント」とは？

2025/05/28 Yu Yamada

Google Engineers



山田 雄 | Yu Yamada

Customer Engineer
Data Analytics Specialist

Google Cloud
Innovators

- データ分析専門のカスタマーエンジニア
- 15 年以上データ分析基盤の開発・運用に従事
- 日本最大級のデータ エンジニア コミュニティ主催
  - datatech-jp
- 好きなプロダクトは **BigQuery** と **Dataplex**

著書

レビュワー

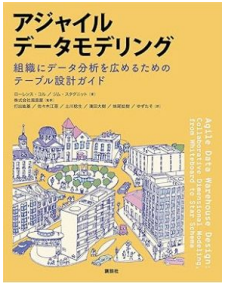


Table of Contents

データ活用の「今」と「未来」 01

Google Cloud の「データ エージェント」とは？ 02

「データ エージェント」がもたらすビジネス変革のシナリオ 03

データ駆動型組織への新たな一歩 04

Table of Contents

データ活用の「今」と「未来」 01

Google Cloud の「データ エージェント」とは？ 02

「データ エージェント」がもたらすビジネス変革のシナリオ 03

データ駆動型組織への新たな一歩 04

The Data Value

raw data から実用的な価値のあるデータへ：速度と複雑性の問題

手作業では時間がかかる

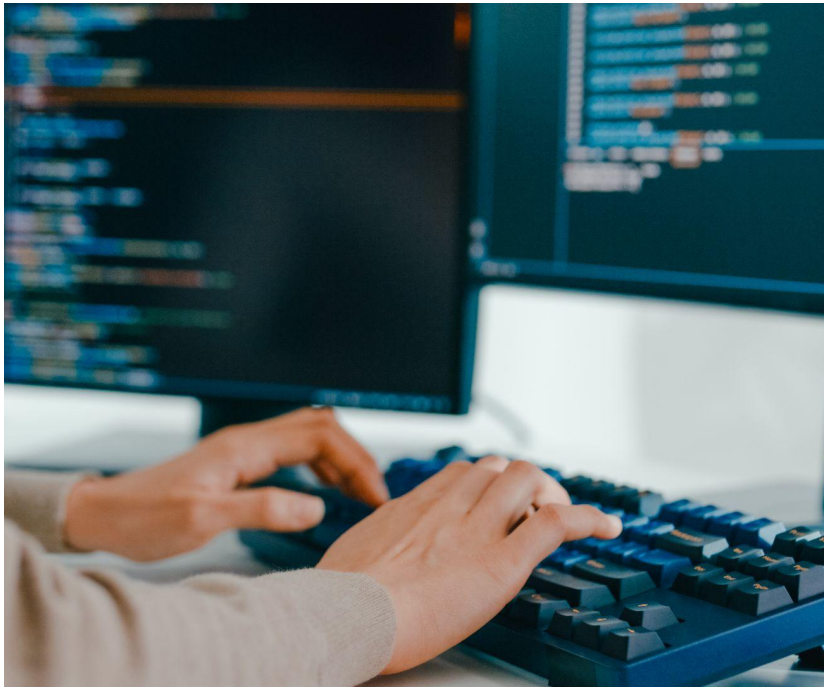
従来の ETLは 手作業でコーディングに時間がかかり時間を消費する

イノベーションが起きづらい

新しいデータを使った実験や競合優位のための高度な分析を迅速に試すことが困難

サイロとマルチモーダル

様々なシステムをつなぎ合わせた分析や、様々なフォーマットの分析が困難



データを取り扱う上での課題

専門家でないとデータから価値を引き出しにくい

データの民主化が進まない

Autonomous Data and AI platform

Powered by Gemini

“サイロ化されたインサイト ” から “統合インテリジェンス ” へ

データ ウェアハウス

80 年代後半 – 2000 年初期

- 構造化データのみ
- 高い所有コスト
- スケール制限

決まった分析 , 限られたデータ

何が起こったかの報告や事前定義された問いへの回答に最適

+

データレイクと データ ウェアハウス

2000 年初期 – 2010 年中期

- 全てのデータタイプ
- データサイエンスの台頭
- スケーラブル & コスト効率

データ種類に対応 , サイロ化が進む

データ サイエンス、探索的分析、ML モデル

=

レイクハウス

2010 年中期 – 2020 年初期

- 統合データ プラットフォーム
- オープンデータ フォーマット
- 統合ガバナンス & セキュリティ

レイクとウェアハウスの いいところどり、オープン

構造化データと非構造化データを組み合わせて、何が起こったのかという問いに答え、ML/AI モデルを構築

自律的 Data & AI platform

2020 年 – 現在

- AIネイティブ
- マルチモーダル
- Agentic

統合 , エージェント 、柔軟性

マルチモーダル データのリアルタイム インテリジェンスと、すべてのデータおよびビジネス プロフェッショナル向けのエージェント型 AI エクスペリエンス

BigQuery

データの取り込みから AI による分析まで、データのライフサイクル全体を支援して、手作業の労力を最小限に抑える

BigQuery と Vertex AI のシームレスな統合により、ML/AI モデルのライフサイクルを迅速に加速し、多様なチームをエンパワー

自律的な data-to-AI platform.

複雑なタスクを自動化し、問題を積極的に特定して解決し、専門のアナリスト、エンジニア、保護者として機能する AI エージェント

マルチモーダル データ、AI 駆動型分析、セマンティック インテリジェンス、LLM コプロセッシング、多様なエンジン、会話型 BI、リアルタイム機能を統合した統合プラットフォーム

Table of Contents

データ活用の「今」と「未来」 01

Google Cloud の「データ エージェント」とは？ 02

「データ エージェント」がもたらすビジネス変革のシナリオ 03

データ駆動型組織への新たな一歩 04

AI Agentとは？

特定の目的を達成するために、環境を認識し、状況を判断し、自律的に行動を実行する　AI システム。

以下のような要素を持ち合わせている

認識（Perception）：まるで人間の目や耳のように、周囲の状況や情報（例：ユーザーからの指示、アクセス可能なデータ、他のシステムからの情報）を理解する能力。

推論・計画（ Reasoning & Planning）：認識した情報に基づいて、目標達成のための最適な方法を考え、どのように行動すべきか、具体的な計画を立てる能力。これは、人間が「どうすればうまくいくか」を考えるのと同じようなプロセスです。

行動（Action）：計画に従って、実際にシステム内で具体的なタスクを実行する能力。例えば、データ分析の実行、レポートの作成、他システムへの連携などです。

記憶・学習（ Memory & Learning）：過去の対話や実行結果を記憶し、そこから学び、次回のパフォーマンスを向上させる能力。これにより、使えば使うほど賢く、パーソナライズされていきます。

自律性（Autonomy）：そして、これが最も重要な点ですが、人間の細かい指示がなくても、自ら状況を判断し、複数のタスクを連携・実行できる能力を指します。

What is an ‘AI agent’?

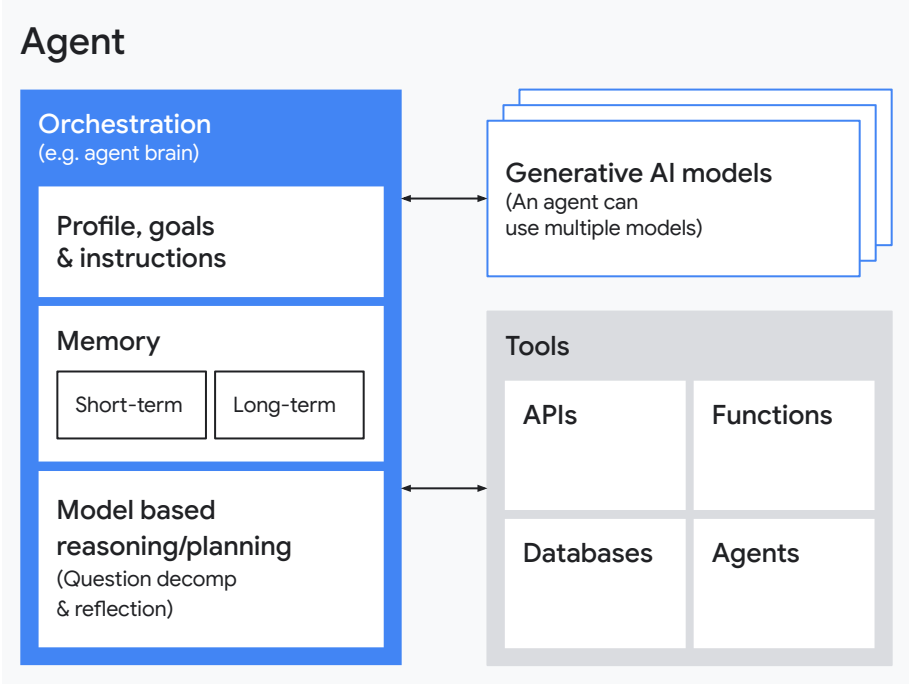
入力情報とツールを活用し、目的のための最適な方法を推論するアプリケーション

Key components

Model:
目標を推論し、計画を決定し、レスポンスを生成するために使用されます。

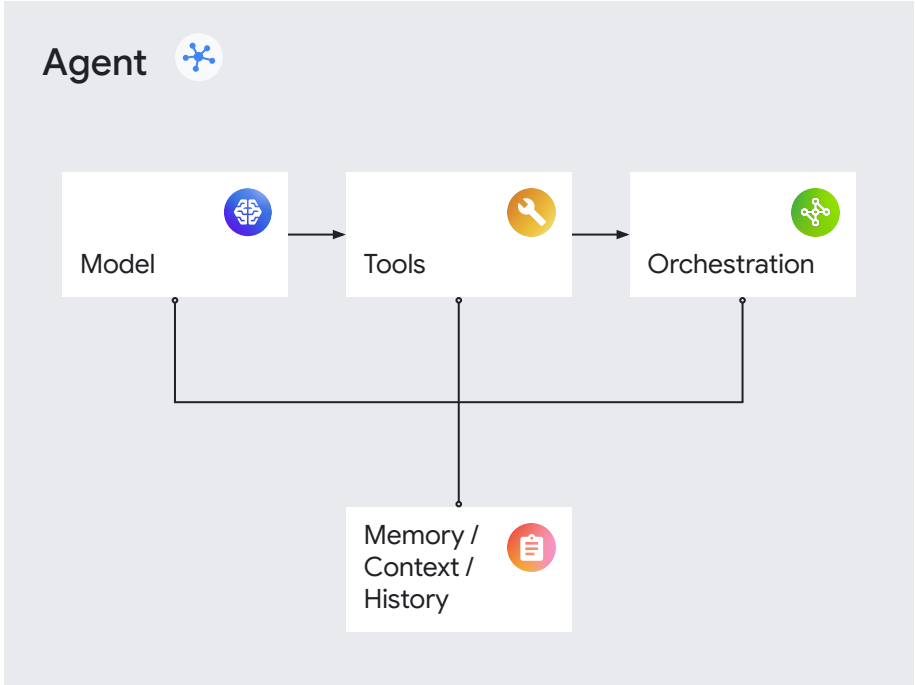
Tools:
他のAPIやサービスを呼び出すことで、データを取得したり、アクションやトランザクションを実行したりします。

Orchestration:
メモリと状態（計画に使用されたアプローチを含む）、ツール、提供されたデータや取得されたデータなどを維持・管理します。



Building blocks of an agent

Model	**The Brain**
Tools	**The Hands**
Orchestration	**The Coordinator**
Memory/ Context / History	

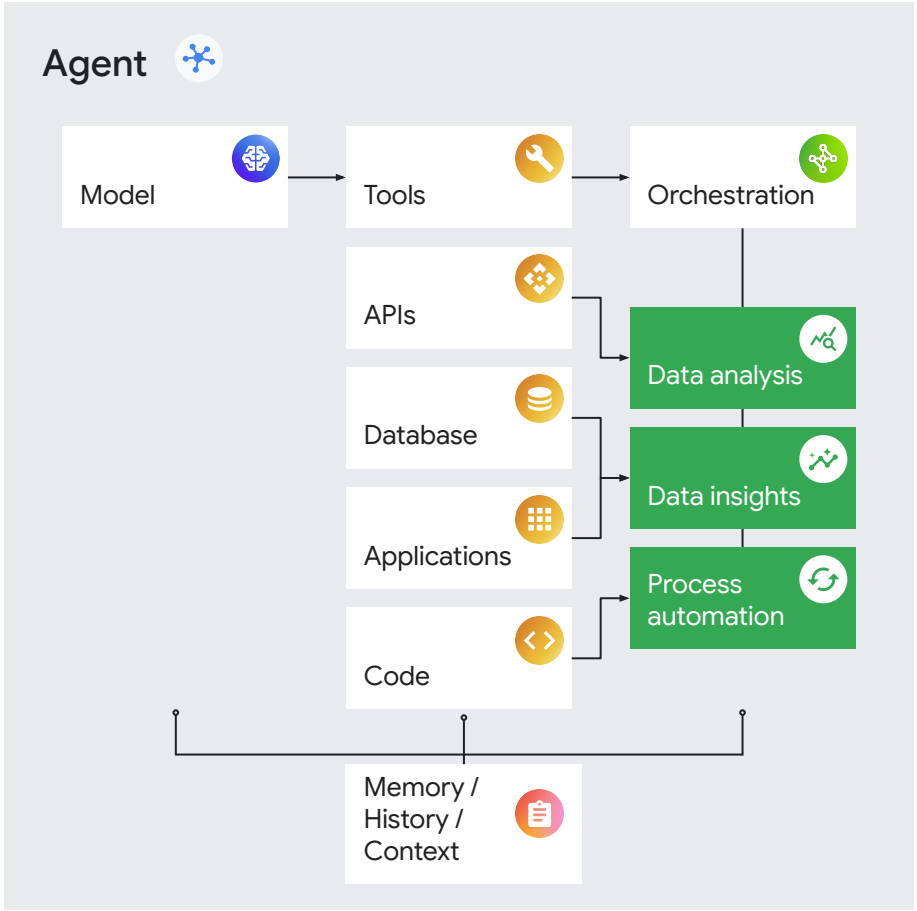


Agents in action: The power of tools

エージェントが利用するツールは一つだけではありません。常に業務に合わせて、最も適切なツールを選択します

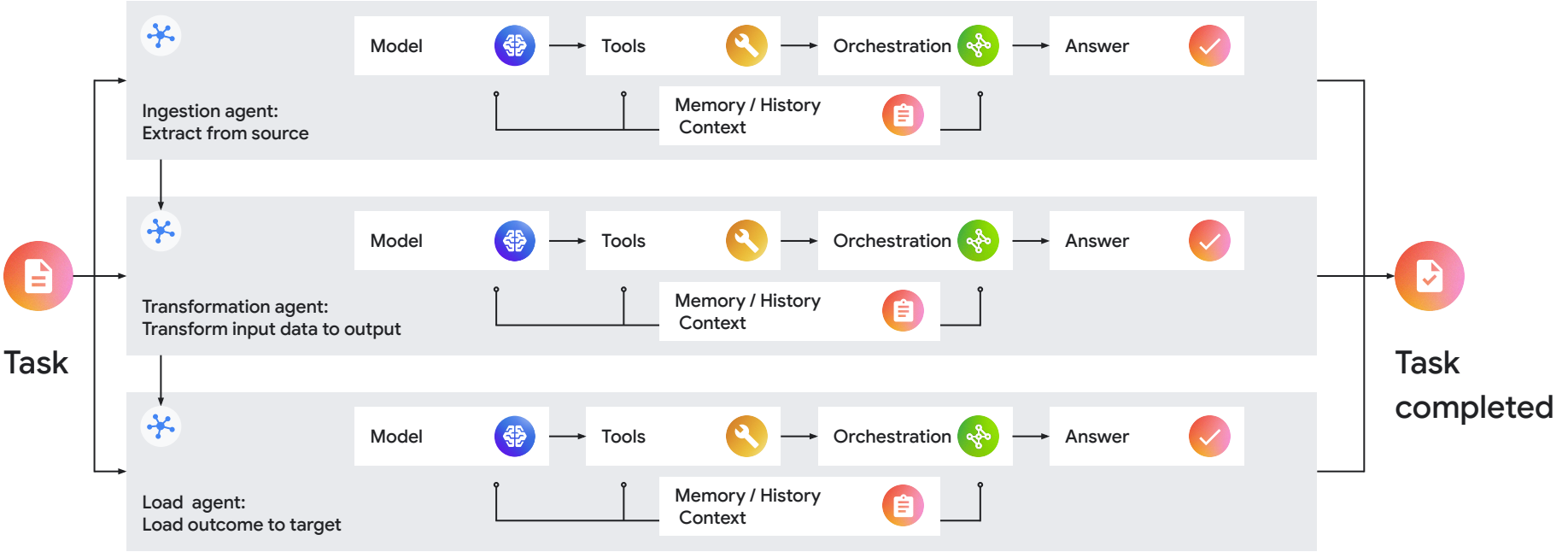
データ分析からプロセス自動化まで、強力な機能を活用可能に

シームレスに連携する、単一の統合型ツールプラットフォーム



From individuals to teams: The multi-agent approach

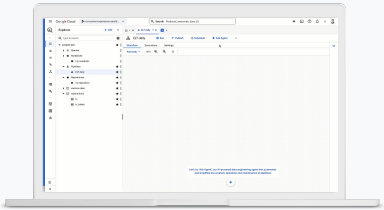
専門スキルを持つエージェントたちがチームを組み、互いの知見を共有しながら連携して、大きな課題を解決



Data Agents

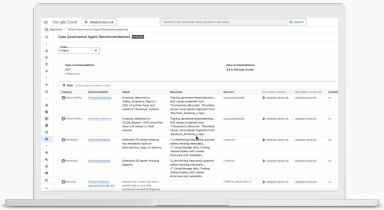
ビジョン　ロードマップ

データ エージェント ファミリー



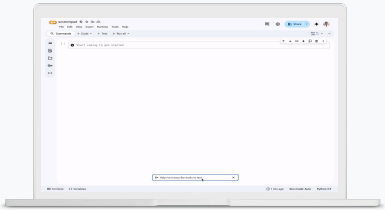
Data Engineering Agent

データ準備、データ取り込み、データ クリーニング、ブロンズ→シルバー→ゴールドの自動化



Data Governance Agent

データ モデリング、メトリック定義、メタデータ生成



Data Science Agent

Colab ノートブック エージェント、探索的データ分析 (EDA)、データ可視化、予測モデリング

Conversational Analytics Agent

Data QnA, フォーキャスト,

マルチモーダル検索

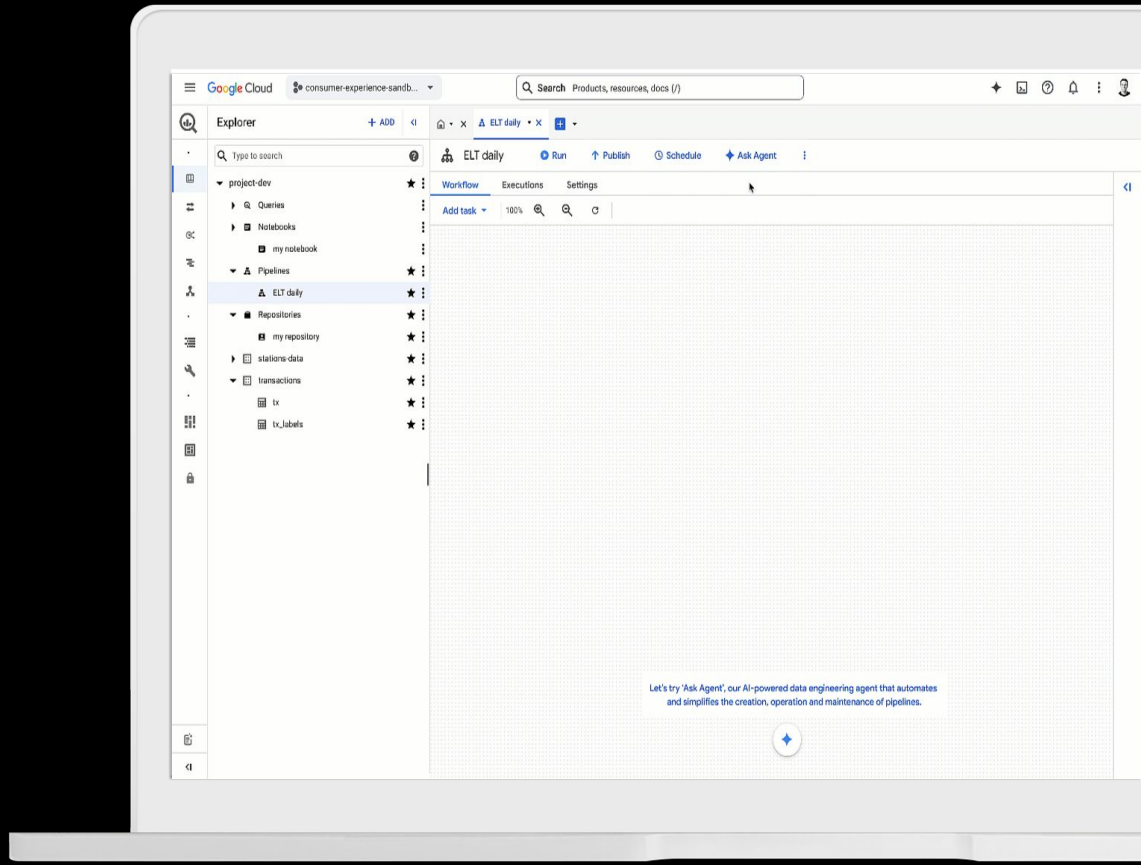
ロードマップ

Data Engineering Agent

自動化されたデータ パイプラインの作成、キュレーション、最適化。BigQuery とのシームレスな連携

- データ パイプラインを大幅に加速
- 全てのデータ関連ステークホルダーを支援
- データ品質とガバナンスを保証
- データ エンジニアリング コストを削減
- ビジネス インサイトを即座に解き放つ

データ エンジニアの工数削減に

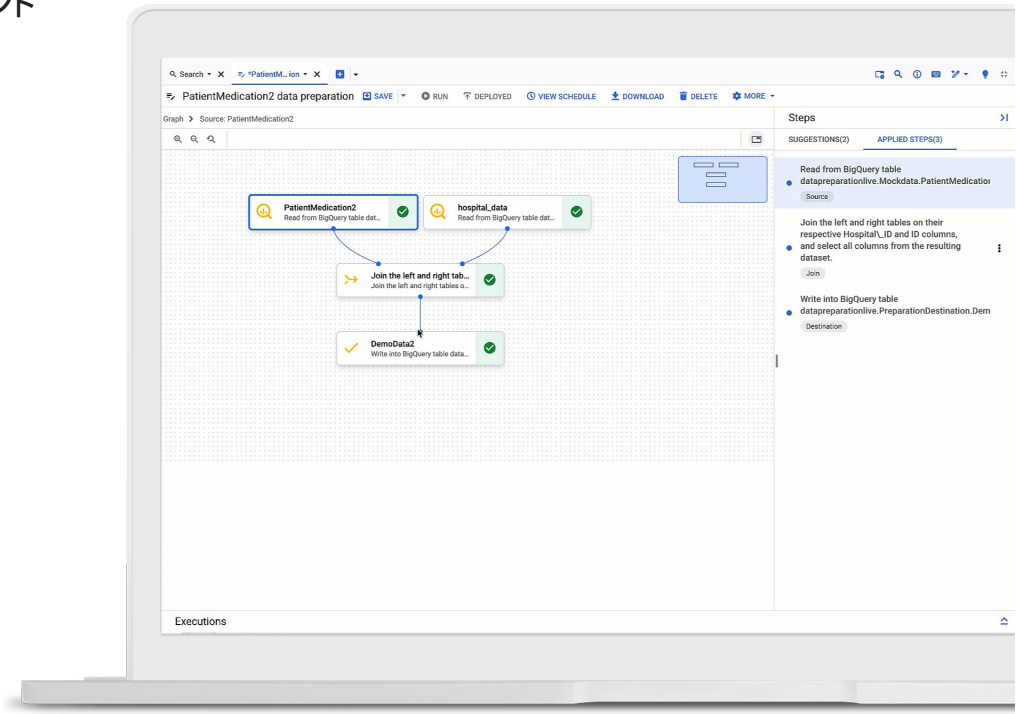


一般提供

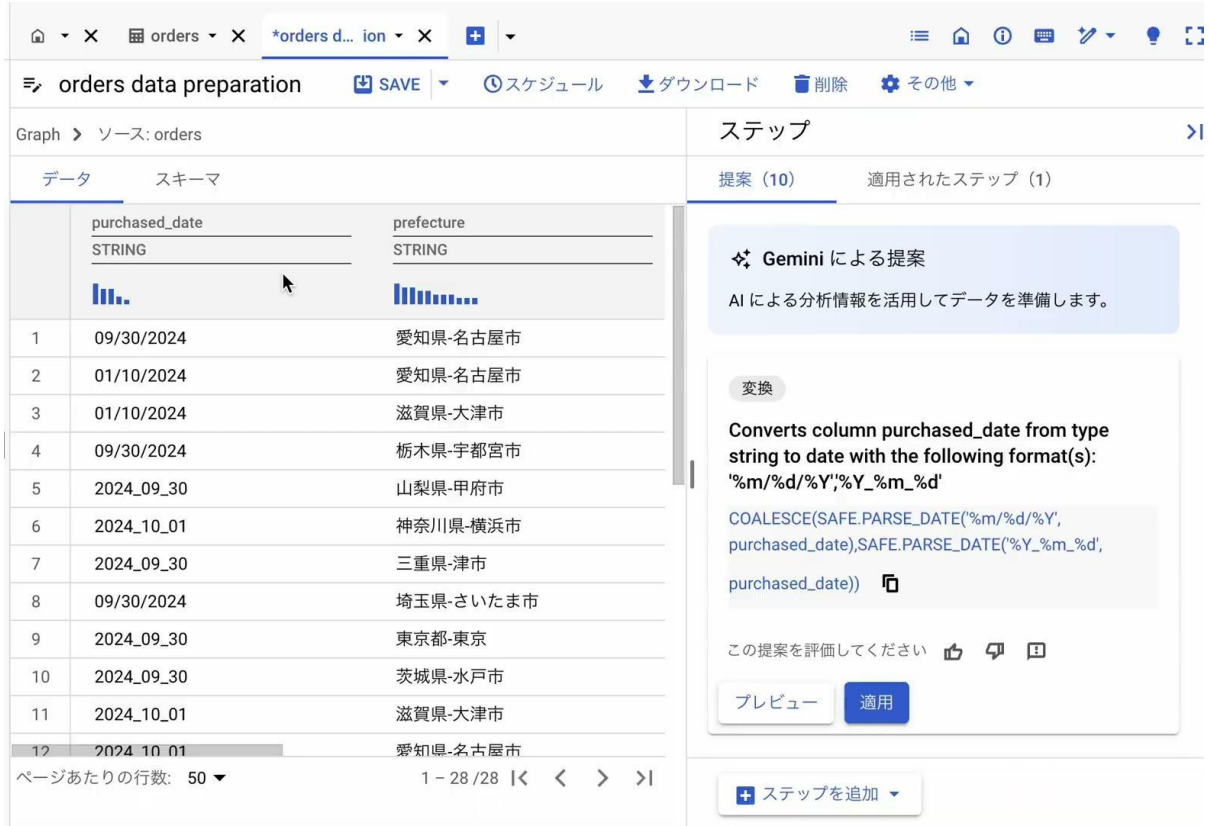
BigQuery Data Preparation

BigQuery Studio 内のデータプリパレーション エージェント

- Gemini の支援により
  SQL データ パイプラインを数分で開発
- コンテキストを考慮した
  AI 生成の変換機能により、データの検出、
  クレンジング、変換を実行可能に
- スキーマ マッチング、
  ビジネスとデータ品質のアサーション生成を
  自動化することで、時間を節約
- GitHub との連携により、BigQuery で
  デプロイとオーケストレーションを実行可能



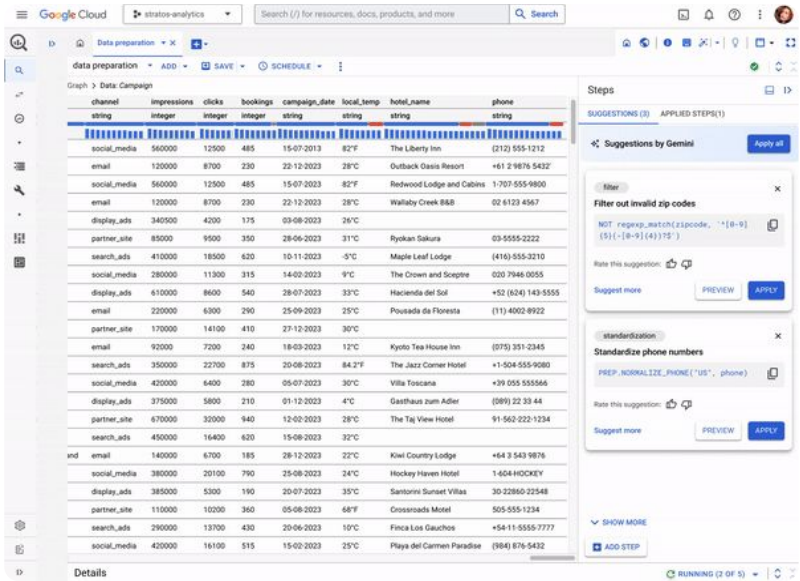
Data Preparation / 注文データをクレンジング



Gemini in BigQuery - AI for everyone

Gemini の主要機能が無料で利用可能に

- SQL コード生成 / 説明 / 補完
- Python コード生成 / コード補完
- Data canvas
- Data preparation



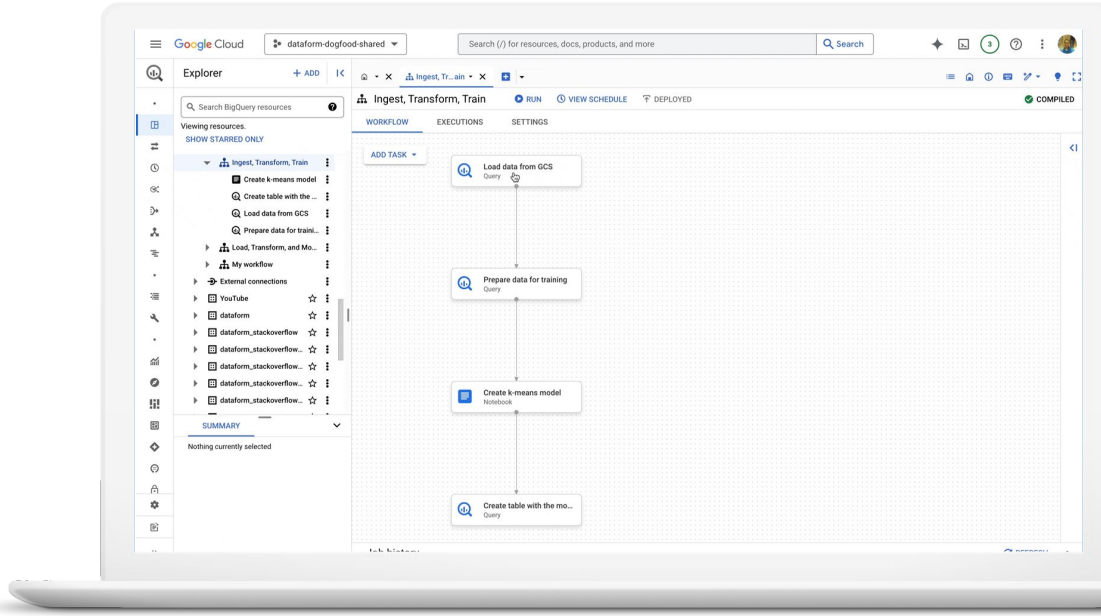
高度なコンテキストとプロアクティブな管理ツールを柔軟な価格帯から選択

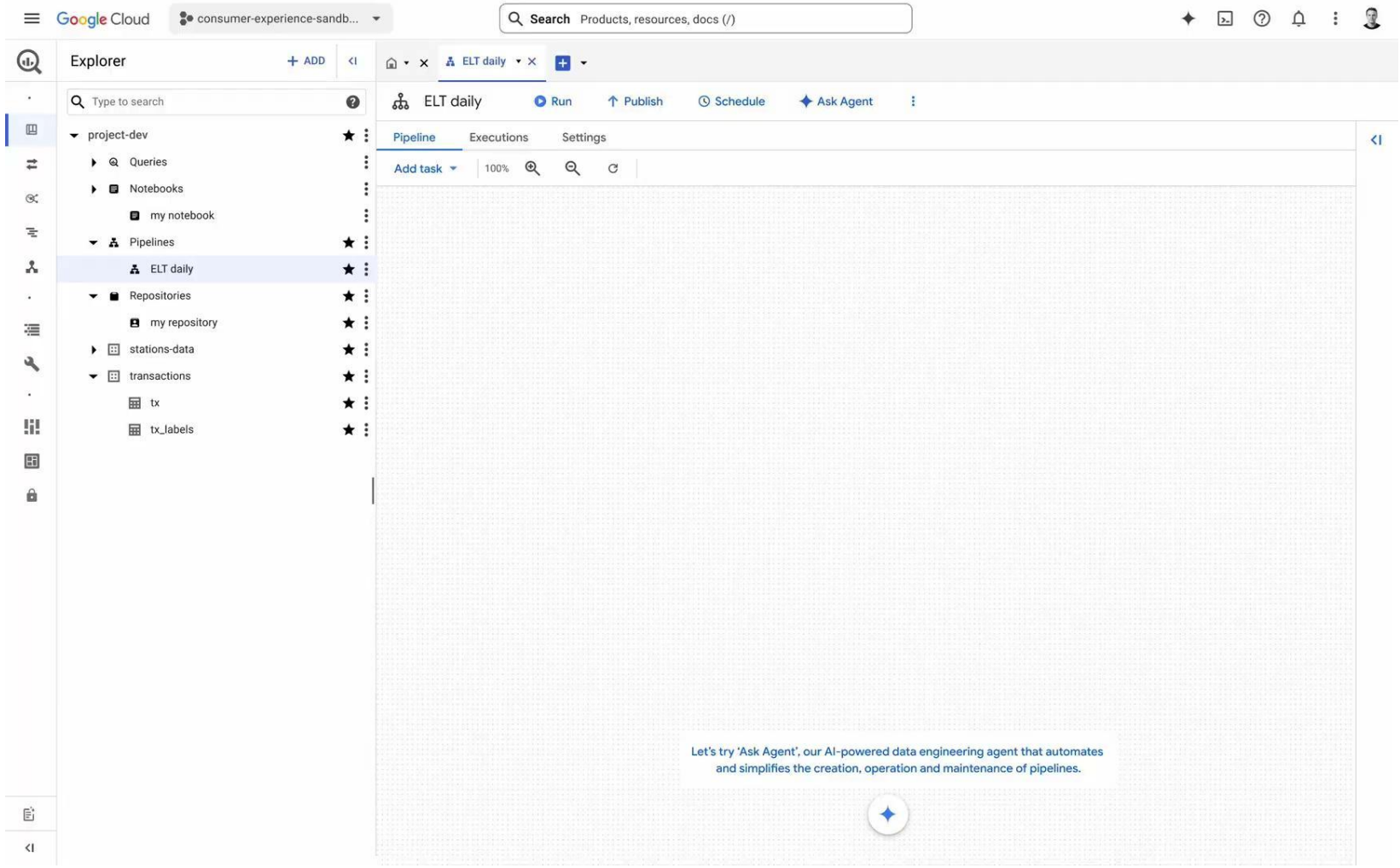
一般提供

BigQuery Pipeline

グラフィカルなパイプラインをBigQuery 上で構成

- ワークフローで SQL スクリプト、Python ノートブック、Data Preparation を定期実行可能
- ノーコードで パイプラインを視覚的に構築
- BigQuery Studio に統合 BigQuery のすべてのスケジュール ジョブを 単一画面から監視
- エンジニアに頼らずワークフローの作成が可能



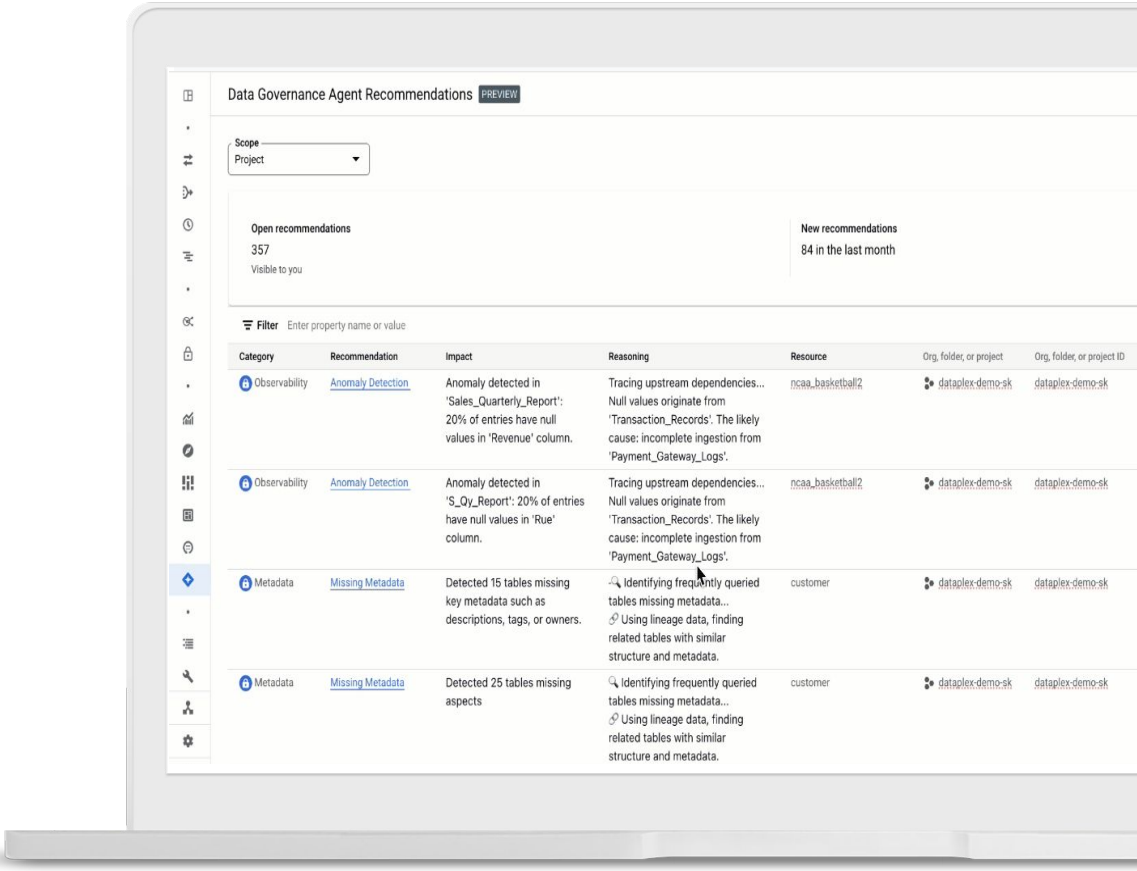

Google Cloud
consumer-experience-sandb...
Search Products, resources, docs (/)
Explorer
+ ADD
Type to search
project-dev
Queries
Notebooks
my notebook
Pipelines
ELT daily
Repositories
my repository
stations-data
transactions
tx
tx_labels
ELT daily
Run
Publish
Schedule
Ask Agent
Pipeline
Executions
Settings
Add task
100%
Let's try 'Ask Agent', our AI-powered data engineering agent that automates and simplifies the creation, operation and maintenance of pipelines.

ビジョン

Data Governance Agent

データ ガバナンス タスクを自動化

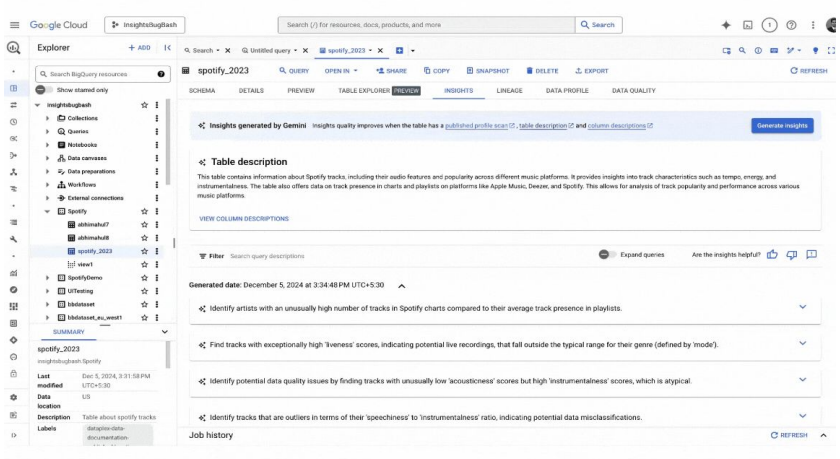
- Proactive Governance: コンプライアンス上の不備が発生するのを待つのではなく、システムがデータ分類、アクセス制御、データリネージ追跡、ポリシー更新を積極的に提案
- Data enrichment: テーブルや列のメタデータを自動的に生成し、拡充
- Semantic grounding for agents: エージェントにとって、統制の取れた信頼できる唯一の情報源（SoT: Source of Truth）
- Knowledge Graph-Powered Insights: ナレッジグラフを利用してデータセット、メタデータ、ポリシー間の関連性を解明（マッピング）し、リアルタイムでガバナンスに関する推奨事項を提供



Experimental ロードマップ

Automated metadata curation

- 自動メタデータ キュレーションは、明確なドキュメントの欠如や、エラーが発生しやすい手動メタデータ作成に起因する課題に対処します。特にデータセットの規模と複雑さが増すにつれて、こうした課題が顕在化します
- BigQuery データ インサイト スキャンは、**テーブルと列の説明**、そしてクエリの推奨事項を**自動生成**します
- ユーザーは必要に応じて説明を確認、編集し、スキーマタブと詳細タブにコピーできます



Experimental　ロードマップ

自動的なデータ異常探知

データの信頼性を向上

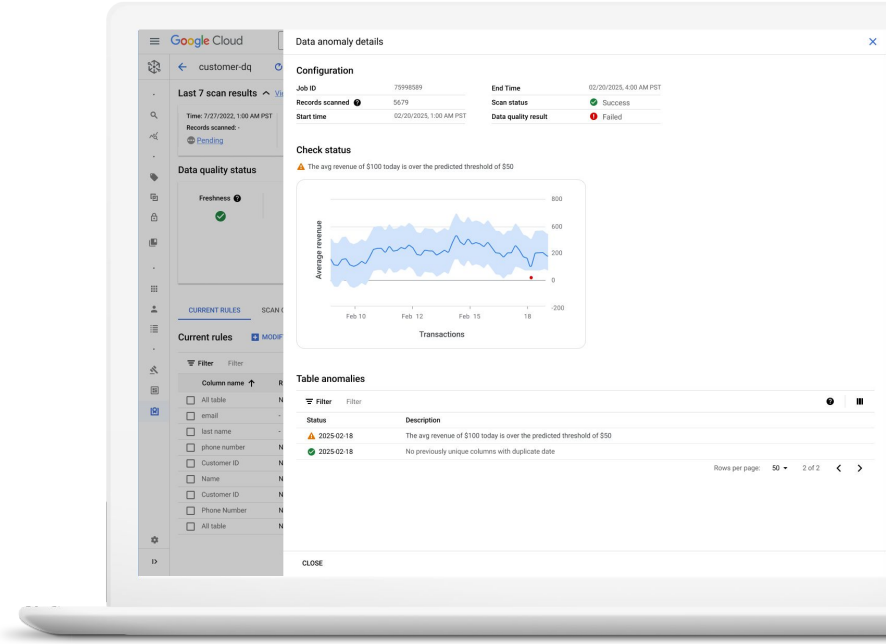
BigQuery の Data Quality 機能で、
BigQuery テーブルとビューのデータの問題を監視して
自動的に検出する自動データ異常検出のサポートを開始

能動的なリスク対処が可能に

異常を特定し、コストのかかるミス、不正行為、
コンプライアンス違反がビジネスに影響する前に防止

AI の力を解き放つ

高度な機械学習アルゴリズムを活用して、
従来のルールベースのシステムでは見逃される
可能性のあるパターンや異常を自動的に検出



ロードマップ

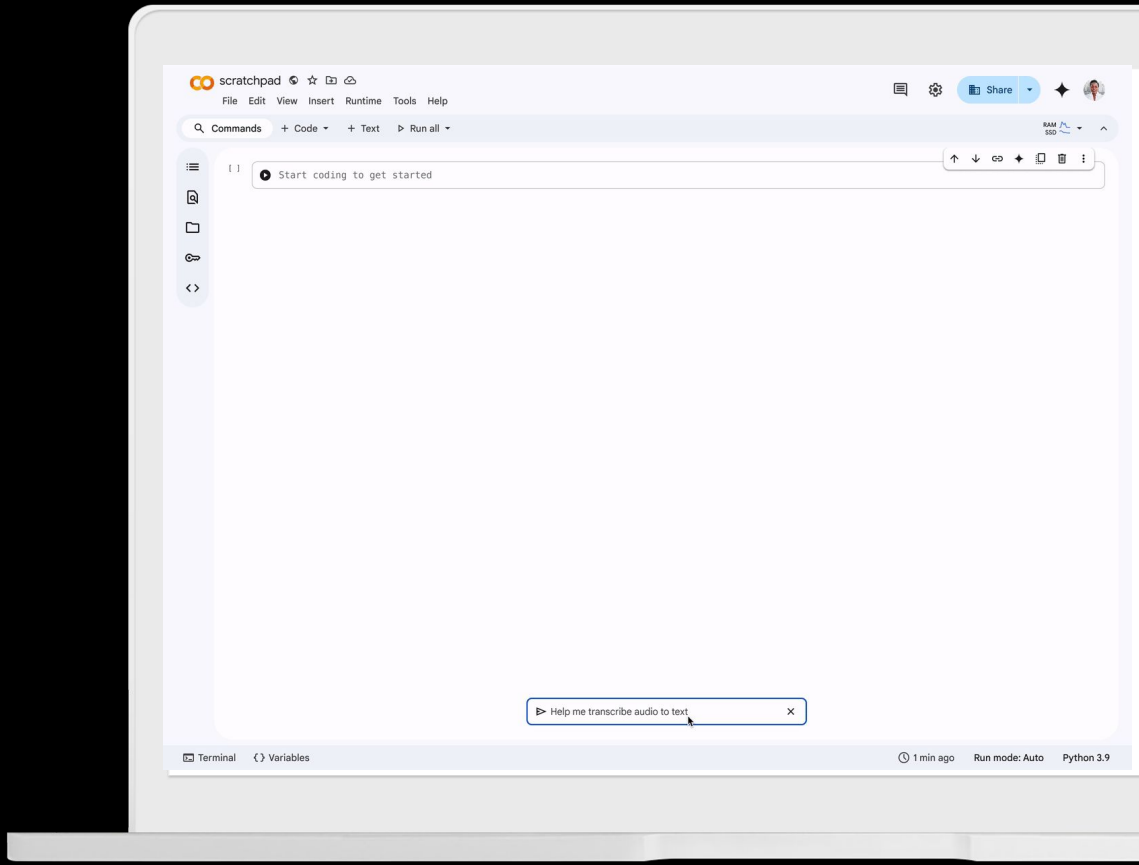
Data Science Agent

データ サイエンスのタスクを自動化する 1st Party エージェント

以下の作業を自動化 ：

- **データ ラングリング** ：データ フレームの自動取り込みとクリーニング
- **データ探索** ：ネイティブの視覚化と自動探索分析
- **AIによる予測** ：Google Research の TimesFM などの最先端のモデルを含む AI を活用した予測
- **モデル評価** ：精度と生産準備状況をより深く理解するためのモデル評価

Callout Agent session



Google Cloud

Proprietary & Confidential

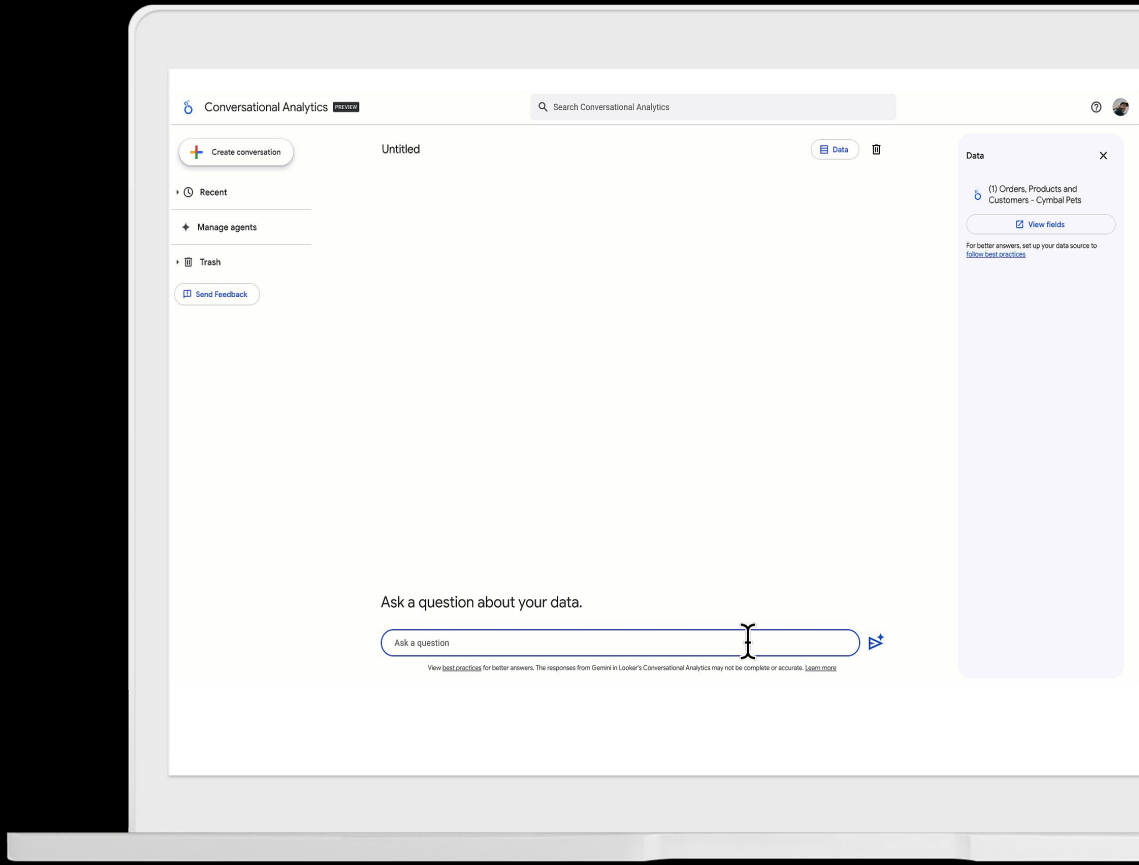
プレビュー

Conversational Analytics Agent

データ分析タスクを自動化する
1st Party エージェント

以下の作業を自動化：

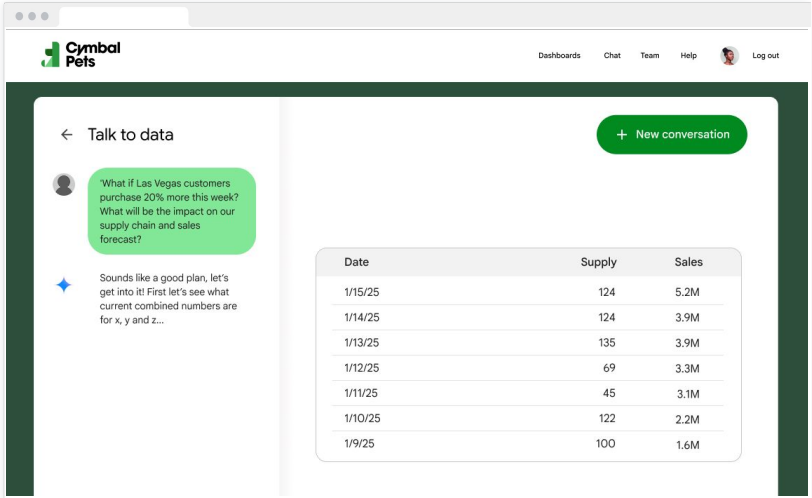
- **分析の民主化：**自然言語ベースのインサイト
- **Custom data agents:** 迅速な導入を可能にするクイックスタートを備えた柔軟な API
- **Direct connect to BigQuery:** BigQuery でデータを直接探索するか、Looker のセマンティック レイヤーと組み合わせる
- **Python の Code interpreter:** 複雑なデータ サイエンス タスク用の Python コードを生成



プライベート プレビュー

Conversational Analytics API

エンタープライズ横断で Data Agent をカスタマイズ & デプロイ



BigQuery

Looker Model

カスタム エージェントを迅速に開発

クイックスタートを備えた
柔軟な API で迅速な導入を実現

分析目的に特化した API

ビジネス インテリジェンス、データ探索、
分析向けにチューニングされたAI サービス

BigQuery へ直接接続可能

BigQuery でデータを直接探索するか、
Looker のセマンティックレイヤと組み合わせる

再利用可能でスケーラブルな セマンティック レイヤー

精度の向上と高度なマルチテーブル分析

**Google Cloud コンソールに入らずに
Slack など普段使い慣れた IF からのデータ分析が可能**

Chat

Search in chat

Active

Looker Data Barista

Chat Home About

Now

Whats the average sales for first time visitors

Looker Data Barista App Now

Average sales for first time visitors is **4.30 USD**. This only includes registered customers!

Now

what questions can you help me with?

Looker Data Barista App Now

This datasource contains information about the state of our Looker Coffee business. Here are some example questions you could ask:

- How many total orders did we serve?
- What is our customers favorite product?
- How are our locations in Berlin performing?

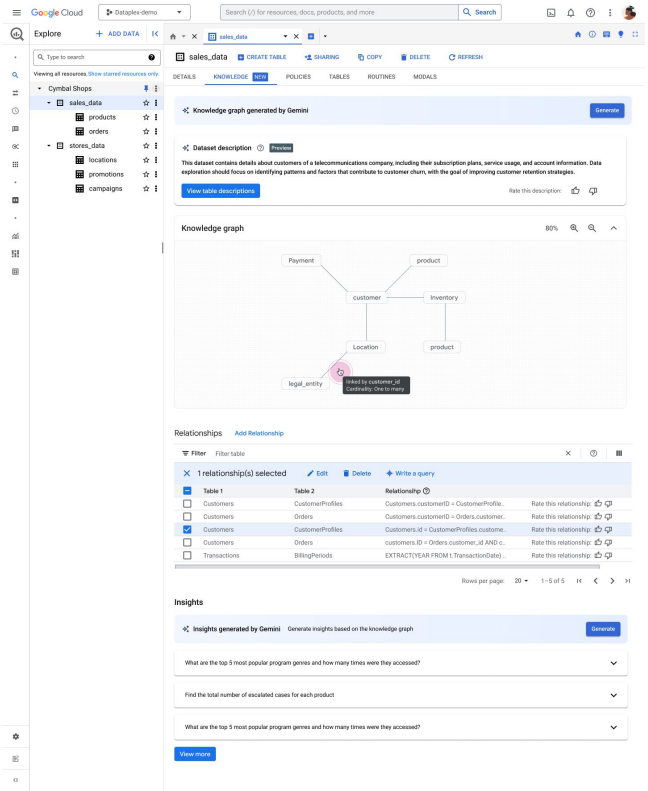
You can also share images with me along with a question.

History is on

Experimental ロードマップ

BigQuery Knowledge Engine

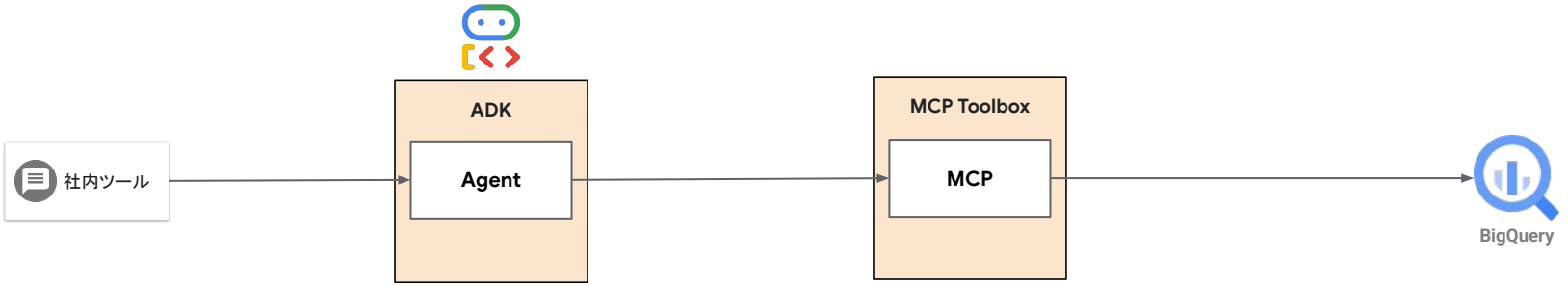
- BigQuery Knowledge Engine は、**Knowledge Graph** の力を活用して、高度なデータインサイトとデータ連携を実現します。
- 多様なデータセットとメタデータを統合することで、ユーザーはテーブル間の**隠れた関係性**を発見し、テーブル間クエリを生成し、より深い分析インサイトを得ることができます

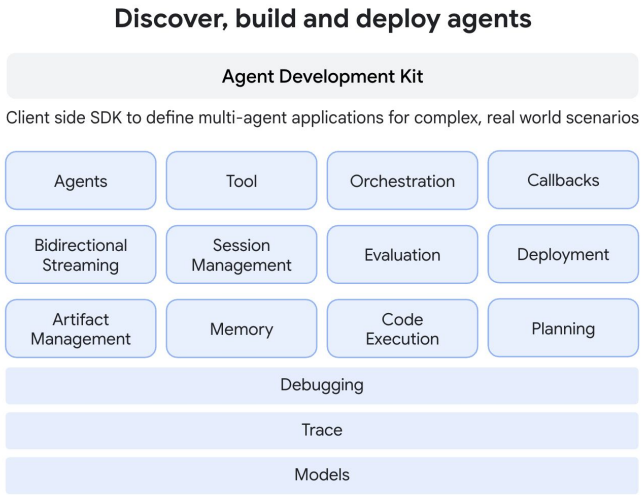


MCP サーバと独自 Agent

社内ツールなどから独自 Agent を使い MCP を経由して BigQuery のデータ操作も可能
セキュアな環境にて構築がしやすい

MCP サーバと自前 Agent





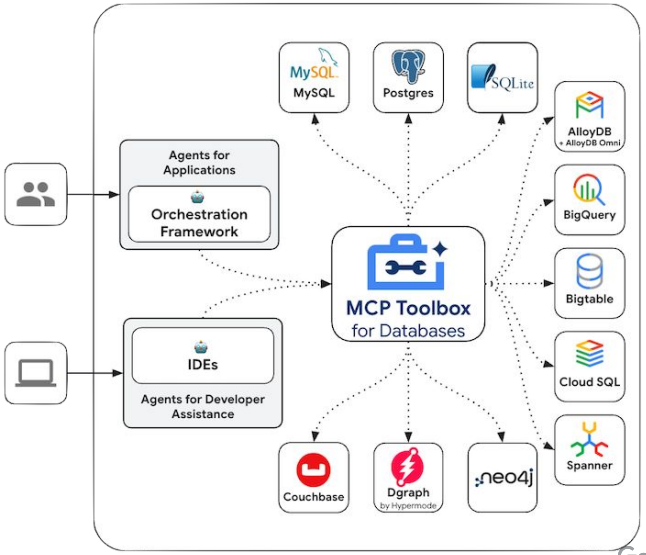


Table of Contents

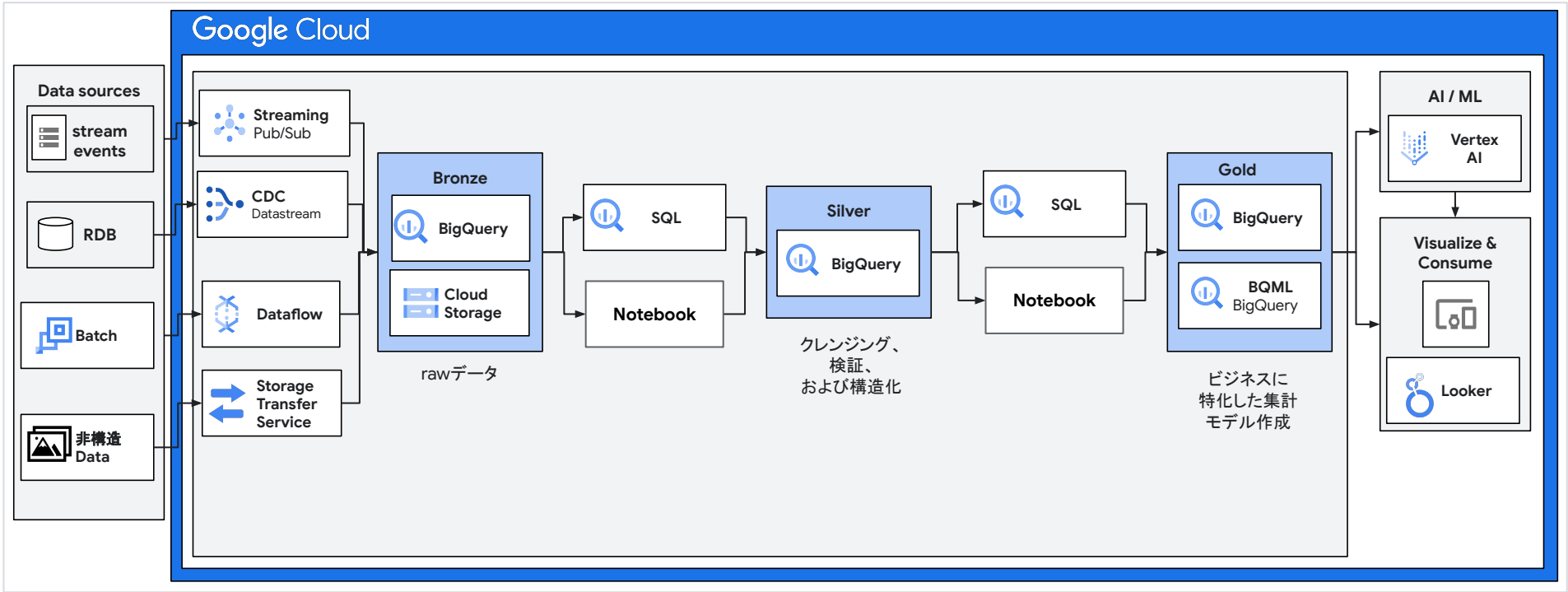
データ活用の「今」と「未来」 01

Google Cloud の「データ エージェント」とは？ 02

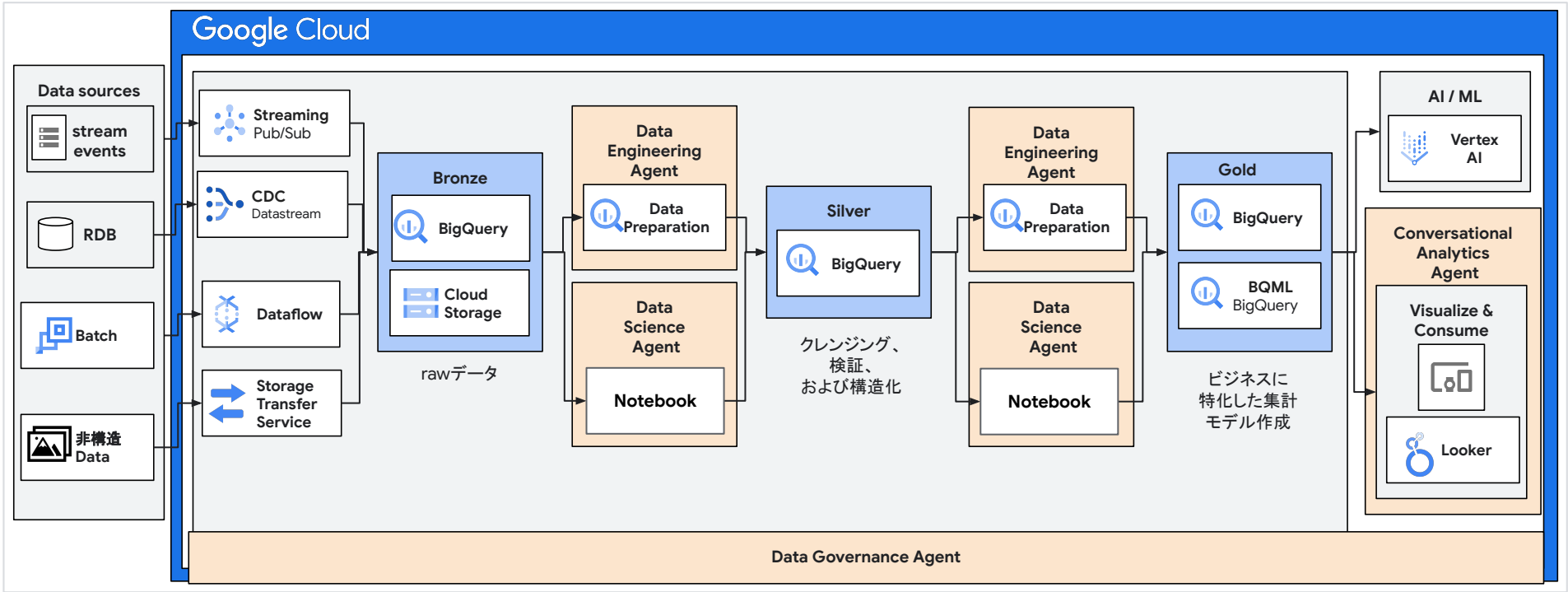
「データ エージェント」がもたらすビジネス変革のシナリオ 03

データ駆動型組織への新たな一歩 04

メダリオン アーキテクチャ

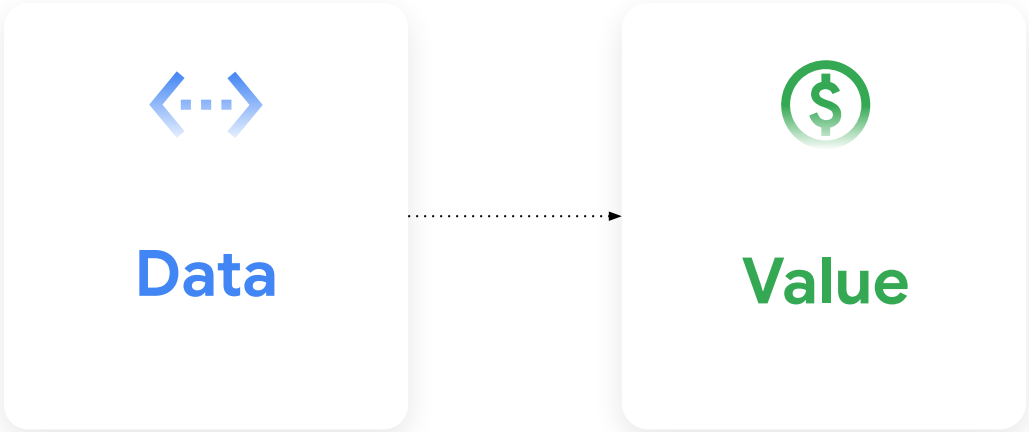


メダリオン アーキテクチャ with Data Agents



The Data Value

The Data Value



部門・役割ごとの未来像

部門・役割ごとの未来像：

- 営業担当者：「**この地域の最新の売上トレンドと、成功しているキャンペーンは？**　」といった質問をデータ エージェントに投げかけ、リアルタイムで顧客データや市場動向を分析し、商談に活かす。
- マーケティング担当者：「**最新のキャンペーンの ROI を前年と比較して教えて。効果が上がっている顧客層は？**」といった自然言語での問い合わせから、瞬時に詳細な分析レポートや改善案を得る。
- 経営層・事業責任者：「**次の四半期の成長ドライバーとなる事業は何か？リスク要因は？**　」といった問いに対して、複数のデータソースを統合したインサイトを迅速に得る。
- データ アナリスト / エンジニア：**ルーチンワークから解放され**、より高度な戦略的分析やデータ基盤構築に注力できる。

データ民主化、**意思決定の加速**、**生産性の向上**、**新たなビジネス価値の創出**

Table of Contents

データ活用の「今」と「未来」 01

Google Cloud の「データ エージェント」とは？ 02

「データ エージェント」がもたらすビジネス変革のシナリオ 03

データ駆動型組織への新たな一歩 04

エンド ツー エンドのデータ分析プラットフォーム

収集

スケーラブルな
データ収集

Pub/Sub

BigQuery Data Transfer Service

Datastream

Storage Transfer Service

処理

信頼性のある
ストリーム データ
パイプライン

Dataflow

Dataproc

蓄積

データレイク
データ ウェアハウス

Cloud Storage

BigQuery
storage

分析

データ ウエアハウス

BigQuery
analysis engine

活用

先進的な分析

Vertex AI

Sheets

Looker

Looker Studio

Cloud Data Fusion

Dataplex（データガバナンス）

Cloud Composer

エンド ツー エンドのデータ分析プラットフォーム

Google Cloud

Google Cloud